Title	Vocal tract dimensional characteristics of professional female singers with different types of singing voices
Author(s)	Yan, N; Xue, SA; Mok, KM
Citation	The 4th International Conference on BioMedical Engineering and Informatics (BMEI 2011), Shanghai, China, 15-17 October 2011. In Proceedings of the 4th BMEI, 2011, p. 1549-1552
Issued Date	2011
URL	http://hdl.handle.net/10722/136296
Rights	International Conference on BioMedical Engineering and Informatics Proceedings. Copyright © IEEE.

Vocal tract dimensional characteristics of professional female singers with different types of singing voices

Nan Yan, Steve An, Xue
Division of Speech and Hearing Sciences
The University of Hong Kong
Hong Kong, China

Mok Ka Man
Division of Speech and Hearing Sciences
The University of Hong Kong
Hong Kong, China

***Abstract*—This study investigated the relationship between the female singers' singing voice classification and their vocal tract length, volume, and vowel formant frequencies. Acoustic pharyngometer (ART) was used to measure the vocal tract length and volume of Twenty-seven sopranos, and Twenty-three mezzo-sopranos. The first three formant frequencies of vowels produced by these singers were also obtained. Results showed that sopranos had shorter oral and vocal tract length, smaller oral and vocal tract volume than mezzo-sopranos significantly. Sopranos had higher first three and averaged first three vowel formant frequencies than mezzo-sopranos generally. These preliminary findings indicated that, besides vocal tract length, vocal tract volume may also affect the formant frequencies of the singers, and thus the classification of different types of singing voices.**

Keywords-Vocal tract; Formant frequency; Singer's type

I. INTRODUCTION

The vocal tract is the cavity between the lips and the glottis. It consists of the oral cavity, the pharyngeal cavity, and the nasal cavity [1]. The vocal tract serves as the resonator in human voice production [2]. Sounds, which be produced by vibrating of the vocal folds, are complex periodic waves. The sounds will be amplified when vocal tract dimensions such as volume and length are changed to particular extents. Along the sound spectrum, certain ranges of frequency will be transmitted to the upper parts of vocal tract but other ranges will be damped [3]. The amplified sounds that are transmitted to the oral and nasal cavity will be shaped by articulators such as tongue, soft palate, etc and heard by listeners accordingly. Thus, the function of the vocal tract as the resonator of human voice production has been explained with the source-filter theory and the tube resonator model [2].

Traditionally, vocal pedagogues divided singers into at least three main singing voice types in female: alto, mezzo-soprano, and soprano [2]. A certain type of singing voice possesses particular attributes, and can thus be distinguished from another type of singing voice. For vocal pedagogues, the voice range and the voice timbre are considered as the two main factors in classifying a singer traditionally. Therefore, there are many classification methods for classifying different singing styles. Titze [2] provided the approximate singing fundamental frequencies of different singing voice types, which showed that the lower pitch produce lower fundamental frequency; while the higher pitch produce higher fundamental frequency. But it can be observed that the voice ranges of different types of singing voices overlap. Besides the voice range, vocal pedagogues also rely on the voice timbre during the singing voice classification. Whereas researchers have interested in identifying the factors that create the differences between the singing voice types: what are the physiology of the production of different vocal registers, and their corresponding frequency ranges [4]; what create the differences between the singing voice types and thus enables vocal pedagogues to differentiate them [5-8]? The research of Cleveland [5] suggested the pitch of the fundamental frequency as the main factor for the differentiation of various singing voices. The results showed that in the classification of singing voice, the pitch of the fundamental frequencies served as the "primary cue"; whereas the formant frequencies served as the "secondary cue". It indicated that the vocal pedagogues mainly depended on the pitch of the fundamental frequencies when classifying the voice stimuli into a particular type of singing voice. However, when asked to classify voice stimuli of the same vowel and pitch sung by tenor singers, baritone singers, and bass singers, the vocal pedagogues had to rely on the "secondary cue", the formant frequencies. It was because the fundamental frequency was very similar as the stimuli were of the same pitch. Cleveland found that the mean value of the first four formant frequencies in tenor singers was higher than that of baritone singers, while bass singers had the lowest mean value of the first four formant frequencies. As formant frequencies are determined by the vocal tract configurations, the findings of Cleveland suggested that the vocal tract characteristics might be a contributing factor of the differentiation of singing voices [5]. However, this study also lacked statistical significance, as only eight participants were included. Therefore, the generalization of the findings was not solid. After this research, some researcher still investigated the correlation between the formant frequencies and the singing types [6, 7]. They also found the singer's formant was relevant to the classification of singing voice and the singer's formant is produced by the configuration of vocal tract. There seems to be some consensus that the classification of singing voice is dependent on both the 1) fundamental frequency, which is mainly affected by the length of the vocal folds, and 2) formant frequencies, which are suggested to be affected by the vocal tract. However, the relationship between the formant frequencies and the vocal

978-1-4244-9352-4/11/$26.00 ©2011 IEEE

tract, and their interaction with singing voice classification is still uncertain.

Dmitriev and Kiselev [8] were the first to address the issue. They measured the vocal tract length of 20 Russian professional singers with lateral x-ray technology. In their research, the length of the vocal tract and the corresponding acoustic characteristics (the high and low singing formants) of 20 professional opera singers were studied. It was found that each type of singing voice had a specific range of lower formant frequency and higher formant frequency (see Table I). It suggested that the professional opera singers used a strictly fixed vocal tract configuration (a certain range of vocal tract length) for all vowels, and in the whole voice range of his/her type of singing voice. Therefore, the interests and awareness of the correlation between singing styles and vocal tract configuration was raised. However, the previous studies focused on the vocal tract length. There was not known whether the volume of the vocal tract also interacted with the singing voice timbre, and thus can be used in the singing voice classification. In addition, the x-ray technology can only provide two-dimensional information of human vocal tract. Therefore, volumetric information cannot be obtained directly. Furthermore, use of x-ray was potential harmful because both the examiner and participants were exposed to radiation for a longer than normal time. Side effects might be resulted.

TABLE I. FREQUENCY OF THE LOW SINGING FORMANT, FREQUENCY OF THE HIGH SINGING FORMANT, AND THE LENGTH OF THE VOCAL TRACT OF DIFFERENT SINGING STYLES [8].

Voice type	Frequency of the low singing formant (Hz)	Frequency of the high singing formant (Hz)	Length of the vocal tract (cm)
High soprano	760-800	3100-3500	15.3-16.3
Soprano	700-760	2800-3100	16.8-18.5
Tenor	600-640	2700-2900	19.0-22.0
Baritone	540-600	2500-2700	21.5-24.0

In the current study, a non-invasive device for 3-dimential vocal tract parameters measure, the acoustic pharyngometer (ART, Eccovision Acoustic Pharyngometer; Sensormedics Corp., Yorba Linda, CA), was used to investigate the relationship between the dimensional characteristics (length and volume) of vocal tract and singing voice classification. The ART makes use of the acoustic reflection (AR) technology, which was originally developed for a diagnostic system of the upper respiratory airway diseases [9]. ART has been used to measure the senile changes in the vocal tract [10]. The ART utilizes acoustic reflection signals to generate graphical representation of the measurement of the cross-sectional area from the glottis to the lips. The changes in the cross-sectional area of the cavity can be calculated by comparing the amplitude and temporal changes of the reflected acoustic pulse and the incident pulse. Therefore, the information of the length and the volume of the vocal tract can both be directly obtained, which makes the ART a more convenient means for vocal tract measure than x-ray technology. The process of ART measure is rapid and non-invasive. Furthermore, the accuracy of AR technology in delineating the changes of length and volume of the human vocal tract has been well validated with CT scan [11] and MRI [12].

The aim of this study is to further investigate the relationship between the singing voice classification and vocal tract length, vocal tract volume, and formant frequencies in female singers. It is hypothesized that both the length and the volume of the vocal tract change systematically with different types of singing voice. It assumed that the singers with higher pitch possess shorter and smaller vocal tracts, and voices of higher formant frequencies.

II. METHODS

A. Participants

Fifty-seven female singers were invited to be the participants in this research, which consisted of 32 sopranos and 25 mezzo-sopranos. Forty participants were recruited from the Choir of Simon K.Y. Lee Hall in the University of Hong Kong (HKU). Their singing styles were confirmed by the music director of Choir, who got Master of Music in Choral Conducting. Seventeen participants were recruited from Opera Hong Kong (OHK) Chorus. The type of singing voice and the professional quality of the participants were confirmed by the chorus master of OHK Chorus, who has more than thirty years of experience in opera performance and instructions. The descriptive statistics of the age, height, and weight of the participants are summarized in Table II. All participants reported no history of craniofacial abnormalities and no upper airway diseases at the time of testing. All the participants passed a pure-tone audiometry test bilaterally at 25dB level at the following frequencies: 250Hz, 500Hz, 1000Hz, 2000Hz, and 4000Hz.

TABLE II. MEAN, STANDARD DEVIATIONS AND RANGES OF THE PARTICIPANTS' AGE, HEIGHT AND WEIGHT

	Age (yrs old)		Height (cm)		Weight (kg)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Sopranos	23.56	4.52	160.18	6.89	52.27	7.59
Mezzo-sopranos	25.24	9.63	162.37	6.53	51.32	6.07

B. procedure

The length and the volume of the vocal tract of the participants were measured by using the ART. Each participant was seated upright on a straight-backed chair facing straight ahead during data-recording. Each participant was instructed to mentally produce the vowel /a/ and breath out air from his/her mouth to the wave tube of ART through the mouthpiece. Then, a corresponding area-distance curve that showed the cross-sectional area of the vocal tract as a function of the distance from the lips to the glottis was plotted. Three measurements were taken to participants and another measurement which measured the cross-sectional area-distance curve of nose-breathing through breathing out air from his/her nose was taken. Thus, a total of four area-distance curves that showed the dimensional parameters of the participant's vocal tract were then obtained.

The participants' speech sounds and singing voices were recorded by Olympus WS-200S stereo voice recorder. The participants produced the vowels /a/, /i/, /u/, /æ/, and /o/, and sang the song "happy birthday". The participants were asked to sing with a well-supported voice with modal register, without any adjustment (e.g. vibrato, portamento, staccato, etc.), and changes in vocal register. All recordings were made with a sampling rate of 48k Hz.

C. Analysis

One of the three cross-sectional area-distance curves of mouth-breathing was chosen for analyzing vocal tract dimensions. The choosing criteria was as follow: 1) the oral pharyngeal juncture (OPJ) of the mouth-breathing curve that matched the OPJ of the nose-breathing curve best, and 2) the curve that was most stable: with least fluctuations in magnitude caused by the changes in airflow [10]. The selected area-distance curve was then divided into the following two regions according to the standard criteria outlined by the manufacturer: 1) the oral region: from the incisors to the anterior edge of the OPJ and 2) the pharyngeal region: from the OPJ to the opening of the glottis. There were six vocal tract dimensions (oral length, oral volume, pharyngeal length, pharyngeal volume, vocal tract length and vocal tract volume) obtained from the selected curves [10]. After the above procedure, the means and standard deviations of the six vocal tract dimensions were calculated.

The participant's speech sounds and singing voices were analyzed by Praat software (Paul Boersma and David Weenink). Among the five spoken vowel sounds, the vowel /a/ sound was chosen for analysis. Whereas three sung vowel sounds were chosen for analysis from the singing voice with the following criteria: 1) the vowel /i/ was extracted from the second vowel of 'happy'; 2) the vowel / æ / was extracted from the first vowel of the same word. The first three formant frequencies (F1, F2 and F3) of the selected vowel sounds were obtained with the auto-tracking function of the Praat software.

Mann-Whiney U test was used to compare the results in vocal tract dimensions and formant frequencies between sopranos and mezzo-sopranos.

III. RESULTS

A. Vocal tract length and volume

The means and the standard deviations of the six measured vocal tract parameters (oral volume, oral length, pharyngeal volume, pharyngeal length, vocal tract volume, and vocal tract length) are summarized in Table III. Significant differences were found between sopranos and mezzo-sopranos in oral length ($U = 229.00$, $z = -2.753$, $p = 0.006$), total vocal tract length ($U = 215.00$, $z = -2.978$, $p = 0.003$), oral volume ($U = 247.00$, $z = -2.460$, $p = 0.014$) and total vocal tract volume ($U = 248.00$, $z = -2.444$, $p = 0.015$). The sopranos were found to have significantly shorter oral length and total vocal tract length, smaller oral volume and total vocal tract volume than mezzo-sopranos significantly; whereas no significant difference was found in pharyngeal length and pharyngeal volume of female singers.

TABLE III. MEANS AND STANDARD DEVIATIONS FOR SIX VOCAL TRACT PARAMETERS OF PARTICIPANTS IN FEMALE SINGERS

Vocal Tract Parameters	Sopranos		Mezzo-sopranos	
	Mean	*SD*	*Mean*	*SD*
OL (cm)[a]	9.87	1.07	9.90	1.51
PL (cm)	8.64	1.55	8.11	1.53
TL (cm)	18.51	1.45	18.01	1.09
OV (ml)	51.78	12.23	56.03	17.78
PV (ml)	25.73	8.13	25.40	7.37
TV (ml)	77.51	15.67	81.43	16.20

a. OL: Oral Length, PL: Pharyngeal length, TL: Total Length, OV: Oral Volume, PV: Pharyngeal Volume, TV: Total Volume

B. First three formant frequencies

The means and the standard deviations of the individual first three formant frequencies and the average of three formant frequencies of the four selected vowels (/a/, /æ/, and /i/) were summarized in Table IV. There were no significant differences founded in F1 of all selected vowels. For F2, significant difference was found in the vowel / æ / ($U = 111.50$, $z = -4.640$, $p = 0.000$), whereas no significant differences were found in the vowel /a/ ($U = 361.00$, $z = -0.627$, $p = 0.531$) and vowel /i/ ($U = 347.50$, $z = -0.844$, $p = 0.399$). For F3, significant difference was found in all considered vowels (/a/ ($U = 157.00$, $z = -3.908$, $p = 0.000$), the vowel /æ/ ($U = 52.00$, $z = -5.596$, $p = 0.000$) and vowel /i/ ($U = 170.00$, $z = -3.699$, $p = 0.000$)). Although only F3 of vowel /a/, F2 and F3 of vowel /e/ and F3 of vowel /i/were shown to have significant differences, the mean values of the F2 and F3 of the three vowels in higher singing style (sopranos) were generally higher than those of lower singing style (mezzo-sopranos). The average of the first three frequencies of the three selected vowels sounds produced by Mezzo-sopranos and Sopranos were also compared. The results were shown in Table IV. Significant differences were founded in all of the three vowels between sopranos and mezzo-sopranos, which were vowel /a/ ($U = 257.00$, $z = -2.300$, $p = 0.021$), vowel /æ/ ($U = 28.00$, $z = -5.982$, $p = 0.000$), and vowel /i/ ($U = 217.00$, $z = -2.943$, $p = 0.003$). The results indicated that higher singing style (sopranos) generally had higher averaged first three vowel formant frequencies than those of lower singing style (mezzo-sopranos).

TABLE IV. MEANS OF THE FIRST THREE FORMANT FREQUENCIES OF THE THREE VOWELS OF PARTICIPANTS IN FEMALE SINGERS

Formant frequencies (Hz)		Sopranos		Mezzo-sopranos	
		M	*SD*	*M*	*SD*
/a/	F1	698.82	90.69	698.93	74.52
	F2	1206.35	119.89	1208.37	101.76
	F3	2650.85	220.25	2669.62	189.42
	Average	1518.67	75.01	1525.64	63.07
/æ/	F1	585.80	109.54	598.65	79.27
	F2	1655.82	153.69	1658.17	121.03
	F3	2496.97	150.15	2458.69	110.69
	Average	1579.53	75.35	1571.84	66.59
/i/	F1	366.14	68.27	396.74	68.66
	F2	1962.86	179.36	1996.68	121.42
	F3	2578.17	171.29	2585.81	139. 65
	Average	1635.72	98.55	1649.74	69.72

IV. Discussion

In the study of Dmitriev & Kiselev, they suggested that there was a correlation between the vocal tract length and the formant frequencies structure: the shorter the vocal tract, the higher the formant frequencies, which were correlated to higher singing voice types; whereas the longer the vocal tract, the lower the formant frequencies, which were correlated to lower singing voice types [8]. For female singers, they found that soprano singers had shorter vocal tract and higher formant frequencies than mezzo-sopranos. In the current study, sopranos had significantly shorter total vocal tract length than those of mezzo-sopranos. Also, the formant frequencies of soprano singers were also found to be higher than those of mezzo-soprano singers. Nonetheless, the vocal tract volume of soprano singers was found to be significantly smaller than that of mezzo-sopranos. Therefore, the findings of the current study have supported their works on their suggestion on vocal tract dimensions determine formant frequencies: higher formant frequencies are correlated with the shorter vocal tract length.

Besides vocal tract length, the vocal tract volume of the singers was also investigated. Results showed that sopranos had smaller oral volume and total vocal tract volume than those of mezzo-sopranos significantly. Furthermore, mezzo-sopranos were also found to have significantly lower formant frequencies. From these observations, it is suggested that, besides vocal tract length, the volume of vocal tract may also affect the formation of formant frequencies, which in turn create the difference in timbre among the singing voice types. This assumption becomes more feasible if the findings of Dmitriev & Kiselev [8] and Erickson [7] are considered together. Dmitriev & Kiselev [8] suggested that singers used a strictly fixed length of vocal tract across their whole voice ranges. Whereas Erickson [7] proposed a timbre transformation in singers as they proceeded from their low notes to high notes. If the vocal tract length is the only factor that creates the differences in timbre, the above two findings would then become contradictory. It is because, if the vocal tract length is the sole factor for the difference in timbre, a timbre transformation becomes impossible if singers use a fixed vocal tract length in their whole voice range. Therefore, other factors that interact with the voice timbre are bound to be present. The current study suggests vocal tract volume to be one of the possible factors that may affect the singing voice timbre. To further investigate the relationship between the vocal tract volume and the singing voice classification, study should be conducted with similar procedure, but with more singing voice types in each gender. More representative findings may result when more singing voice types can be compared. The presence of contribution of vocal tract volume to singing voice classification may also be identified in female singers.

V. Conclusion

Two main findings were found in this study. First, in female singers, sopranos posses the shorter oral length and total vocal tract length, smaller oral volume and total vocal tract volume than mezzo-sopranos. Second, in female singers, sopranos had generally higher first three vowel formant frequencies and averaged first three vowel formant frequencies than mezzo-sopranos.

The one limitation of current study is that the singing voice types of participants were classified by different vocal professionals. The classification criteria of the vocal professionals may vary from each other which results in discrepancy in the classifications of singing voice. This discrepancy then decreases the validity of the study. To further investigate the relationship between singing voice types and vocal tract dimensions, one vocal professional instead of several ones should be invited to classify all participants so as to unify the classification criteria and increase the validity of study. In addition, Future studies should involve equal numbers of singers in all the singing categories. Male singers and altos for female could also be recruited as participants for further delineating the effect of singing voice types on vocal tract dimensions, length and volume. To further understand the relationship between the vocal tract characteristics and singing voice classification, studying the relationship between the timbre transformation and its corresponding changes in vocal tract may be useful. As Erickson [7] suggested, singers of a certain singing voice type share a similar pattern of timbre transformation. Therefore, it is expected that each singing voice type may demonstrate a characteristic timbre transformation, and thus a specific pattern of vocal tract configuration. To further investigate this, devices such as electromagnetic articulography could be used that can obtain dynamic changes of the vocal tract during singing performance.

References

[1] J. A. Seikel, D. W. King, and D. G. Drumright, Anatomy & physiology for speech, language, and hearing, 2nd ed., San Diego: Singular, 2000.

[2] I.E. Titze, Principles of voice production. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[3] A. W., Goodwin, "An acoustic study of individual voices in choral blend", J. Res. Music. Educ., 1980, 28, pp. 119-128.

[4] D.G. Miller, "Registers in singing: empirical and systematic studies in the theory of the singing voice," Dissertation, Groningen: University of Groningen, 2000.

[5] T.F. Cleveland, "Acoustic properties of voice timbre types and their influence on voice classification," J. Acoust. Soc. Am., 1977, 61, pp.1622-1629.

[6] G. Berndtsson and J. Sundberg, "Perceptual significance of center frequency of the singer's formant," Scand. J. Logopedics and Phoniatrics, 1995, 20, pp.35-41.

[7] M.L. Erickson, S. Perry and S. Handel, "Discrimination functions: can they be used to classify singing voices?" J. Voice, 2001, 15, 492-502.

[8] L. Dmitriev and A. Kiselev, "Relationship between the formant structure of different types of singing voices and the dimension of supraglottal cavities," Folia Phoniat, 1979, 31, pp. 238-241.

[9] J.P. Corey, A. Gungor, R. Nelson, J. Fredberg, V. Lai, "A comparison of the nasal cross-sectional areas and volumes obtained with acoustic rhinometry and magnetic resonance imaging". Otolaryngol. Head Neck Surg. 1997, 117, 349-354.

[10] S. Xue and JP. Hao, "Gender differences in senile changes of human vocal tract and the acoustic correlates of speech production," J. Speech Lang. Hearing. Res., 2003, 46, 689-701.

[11] A. D. D'Urzo, I. Rubinstein, V. G. Lawson, K.P. Vassal, A. S. Rebuck, A. S. Slutsky, et. al., "Comparison of glottic areas meauresd by scoustic reflection v.s. computerized tomography," J. Appl. Physiol. ,1988, 127, 221-226.

[12] J. Dang, K. Honda and H. Suzuki, "Morphological and acoustical analysis of the nasal and the paranasal cavities," J. Acoust. Soc. Am., 1994, 96, 2088-2100.